



# GRANULAR DATA FICHES

**BENCHMARK OF PERFORMANCE AND  
COSTS ENABLING IDENTIFICATION OF  
VIABLE OPTIONS GOING FORWARD**

## Review & Design

Maite Iglesias (AEIDL), Michael Kull (LUKE), Carla Lostrangio, Merveille Ntabuhashe (both AEIDL).



Funded by  
the European Union



UK Research  
and Innovation

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them. UK participants in the GRANULAR project are supported by UKRI- Grant numbers 10039965 (James Hutton Institute) and 10041831 (University of Southampton).





GRANULAR

DATA FICHES

# ACCESSIBILITY

BENCHMARK OF PERFORMANCE AND COSTS ENABLING  
IDENTIFICATION OF VIABLE OPTIONS GOING FORWARD





GRANULAR

DATA FICHE

# Routing engines based on OpenStreetMap

<https://www.openstreetmap.org/>

Ronan Ysebaert & Marianne Guérois (CNRS)

## INDICATOR CLASS: Accessibility



### Data

Class: Crowdsourcing – Routing engines  
Type: Origins – Destinations matrices / itineraries (lines) / Isochrones  
License: Data Commons Open Database license



### Spatial

Resolution: Itineraries from origin/destination points  
Extent: World  
Granularity: Roads included in the OSM database and its attributes (maxspeed).  
Coordinate System: WGS 1984 (EPSG: 4326)



### Temporal

Resolution: Up-to-date (depend on OSM)  
Frequency: Continuously updated

## DESCRIPTION

Several **open-source routing engines** based on the OpenStreetMap (OSM) database are available to compute travel time matrices between origins and destinations (OSRM, Valhalla, Grasshopper, Openrouteservice, pgRouting). All these engines use the **OSM road network**. It requires as input a set of origin and destination points, whereas outputs are usually routes, trips, isochrones and travel distance matrices (travel time or kilometeric distance).

Routing engines support profiles, representing **routing behaviour** for different transport modes like car, bike and foot. For the car profile, the OSM tag max speed is used to calculate the distance and travel time between two points. If the maximum speed information is missing, which is frequently the case (only 7.4 % of all road elements stored this value in 2019), predefined speed limits for each country are applied.

Bike profile avoids roads without bicycle access. Some routing engines, such as **Valhalla**, provides also **digital elevation model** (DEM) data as the result of the query. Penalties are applied to roads based on elevation change and grade. **Multimodal models** (using several transport profiles for the same route) are not implemented at the moment, but associated documentation of some routing engines (Valhalla, namely) specifies that the implementation of the feature is strongly envisaged for the future.



## USER COSTS

Free & open Access.

## DATA INFRASTRUCTURE



**Infrastructure needed:** A modern computer with 32 GB RAM or more.

**Data repositories/storage needed:** Routing outputs may be usefully stored (O/D matrices) in a csv format.  
Size depending on the number of O/D points considered.



## DATA GOVERNANCE & MANAGEMENT

**Staff costs:** Advanced computer/GIS skills (Docker). Advanced knowledge of the OSM database.

**Data analysis needed:** No. It requires relevant Origin/Destination points as input.

**Quality assurance:** The road network available in OpenStreetMap can be considered of good quality. A completeness analysis led by Barrington-Leigh (2019) based on satellite imagery, multilevel regression and postratification model argued that globally, OSM is 83% complete, and more than 40% of countries-including several in the developing world have a fully mapped street network, concluding that in many places, researchers and policymakers can rely on the completeness of OSM, or will soon be able to do so. Outputs of these routing engines deliver theoretical travel time, without traffic congestion.

The choice of the routing engine depends on several aspects: the activity of the project, hardware requirement, features of the routing engine (turn restrictions, elevation, time awareness), costing options (influence of the route finding with many custom factors/penalties/costs. e.g. avoid highways). For the needs of the GRANULAR project, requiring cover all of Europe (heavy calculations) and several transport profiles, OSRM and Valhalla routing engines appear the most suitable solutions.

## DATA DOCUMENTATION AND/OR ASSISTANCE IN DATA USE



[Luxen, 2011, Real-time routing with OpenStreetMap data](#)  
[OSRM github repository](#)  
[Valhalla github repository](#)

[Giraud, 2023, How to Build a European-Wide OSRM Server on a Desktop Computer](#)  
[Valhallr: A tidy Interface to the Valhalla Routing Engine](#)

### How to cite

©OpenStreetMap contributors





GRANULAR

DATA FICHES

# AGRICULTURE

BENCHMARK OF PERFORMANCE AND COSTS ENABLING  
IDENTIFICATION OF VIABLE OPTIONS GOING FORWARD





# Carbon budget in the EU agricultural soils

<https://esdac.jrc.ec.europa.eu/content/carbon-budget-eu-agricultural-soils>

Tristan Berchoux (IAMM)

## INDICATOR CLASS: Regional



### Data

Class: Model

Type: Raster

License: under NO CIRCUMSTANCES are these data passed to third parties. They can be used for any purpose, including commercial gain.

DOI: 10.1111/ejss.13315.



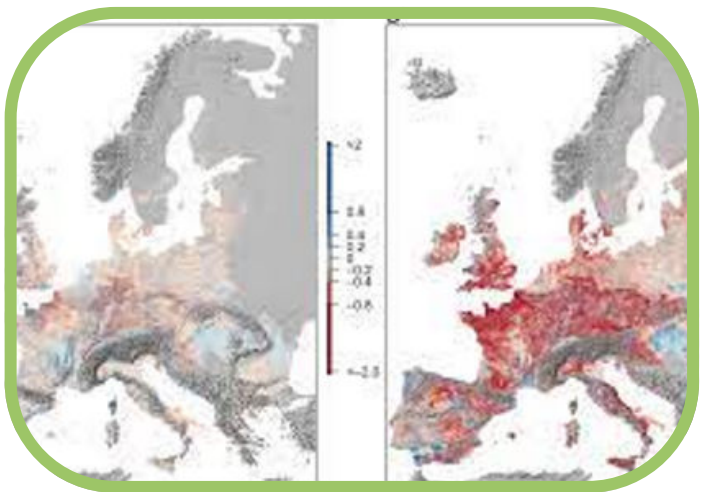
### Spatial

Resolution: 1 X 1 km

Extent: EU

Granularity: Gridded

Coordinate System: ETRS\_1989\_LAEA\_L52\_M10



### Temporal

Resolution: Yearly

Extent: 2016-2100

## DESCRIPTION

Soil plays a significant environmental role in balancing the climate as it currently acts as a carbon sink, sequestering CO<sub>2</sub> from the atmosphere into soil organic carbon. This dataset tracks the possible **transformations of the organic carbon across the landscape** by using a **biogeochemistry-erosion model** to quantify the impact of future climate on the carbon cycle.

Considering all the additional feedback and C fluxes due to displacement by erosion, it is estimated that there will be a net source of 0.92 to 10.1 Tg C year<sup>-1</sup> from agricultural soils in the European Union to the atmosphere over the period 2016–2100. These ranges represented a weaker and stronger C source compared to a simulation without erosion (1.8 Tg C year<sup>-1</sup>), respectively, and were dependent on the erosion-driven C loss parameterization, which is still very uncertain.





## USER COSTS

Free of charge & Access limited (request form).

## DATA INFRASTRUCTURE



**Infrastructure needed:** GIS (QGIS for example) or software environment for statistical computing (R).

**Data repositories/storage needed:** Datasets are hosted on ESDAC and made available for free.



## DATA GOVERNANCE & MANAGEMENT

**Staff costs:** Data can be easily downloaded. Its use requires a certain expertise in the domain of geographic information and the structure of LPIS data (low-level GIS Technician ~25,000€ + charges per year).

**Data analysis needed:** None for simple visualisation, it requires running an R code (supplied) to run simulations.

**Quality assurance:** Data are model-based and might contain errors. Any error or omission should be noted and reported to the Joint Research Centre.

## DATA DOCUMENTATION AND/OR ASSISTANCE IN DATA USE



<https://esdac.jrc.ec.europa.eu/content/carbon-budget-eu-agricultural-soils>.

### How to cite

Dataset: Lugato, E., Smith, P., Borrelli, P., Panagos, P., Ballabio, C., Orgiazzi, A., Fernandez-Ugalde, O., Montanarella, L., Jones, A. 2018. Soil erosion is unlikely to drive a future carbon sink in Europe. Science Advances. 4, eaau3523.



# GRANULAR

## DATA FICHE

### Cover Crops across Europe

<https://esdac.jrc.ec.europa.eu/content/cover-crops-across-europe>

Tristan Berchoux (CIHEAM IAM Montpellier)

#### INDICATOR CLASS: Regional



#### Data

Class: Satellite + Survey

Type: Raster

License: N/A

DOI:

<https://doi.org/10.1016/j.scitotenv.2023.162300>



#### Spatial

Resolution: 100 m

Extent: EU + UK

Granularity: Gridded

Coordinate System: EPSG:3035



#### Temporal

Resolution: Yearly

Extent: 2016

#### DESCRIPTION

Despite the growing importance given to **cover crops** as a **sustainable agricultural practice**, the availability of spatial data about them is scarce. The best information available is regionally aggregated survey data, which, although indicative, hinders the development of spatially accurate studies. Using a disaggregation model, this dataset combines **satellite data** (Sentinel-1) with **aggregated survey data** to generate a high-resolution map of cover crops for Europe and the United Kingdom for the reference year of 2016.

The map was validated with parcel-level data in France, where overall good results were found, with regional variations. This dataset includes the **median, standard deviation, 5th and 95th percentiles of the predicted cover crop fraction**. Such a fraction is assumed to vary between 0% (i.e., no cover crops in the pixel) to 100% (i.e., pixel completely covered by cover crops). The values in the raster files are multiplied by 100 to facilitate compression. For instance, a raster value of 1234 corresponds to a 12.34% cover crop fraction. The dataset aims to replace the practice commonly found in the literature of assigning aggregated values to agricultural pixels by randomly sampling them in space. It can be useful for researchers and practitioners requiring spatially explicit knowledge of cover.





## USER COSTS

Free & open Access (Registration needed).

## DATA INFRASTRUCTURE



**Infrastructure needed:** GIS (QGIS for example) or software environment for statistical computing (R).

**Data repositories/storage needed:** Datasets are hosted on ESDAC and made available for free upon request.



## DATA GOVERNANCE & MANAGEMENT

**Staff costs:** Data can be easily downloaded after completing the request form. Its use requires a certain expertise in the world of geographic information (low-level GIS Technician ~25,000€ + charges per year).

**Data analysis needed:** None for simple visualisation.

**Quality assurance:** Data has good results (validation with parcel-level data in France) but no other details are given. No layer of quality assessment if given.

## DATA DOCUMENTATION AND/OR ASSISTANCE IN DATA USE



<https://esdac.jrc.ec.europa.eu/content/cover-crops-across-europe>

### How to cite

Dataset: Fendrich, A. N., Matthews, F., Van Eynde, E., Carozzi, M., Li, Z., d'Andrimont, R., Lugato, E., Martin, P., Ciais, P., Panagos, P., 2023. From regional to parcel scale: A high-resolution map of cover crops across Europe combining satellite data with statistical surveys. Science of the Total Environment, pp.162300. <https://doi.org/10.1016/j.scitotenv.2023.162300>



# GRANULAR

## DATA FICHE

### EuroCrops

<https://github.com/maja601/EuroCrops>

Tristan Berchoux (IAMM)

## INDICATOR CLASS: Regional



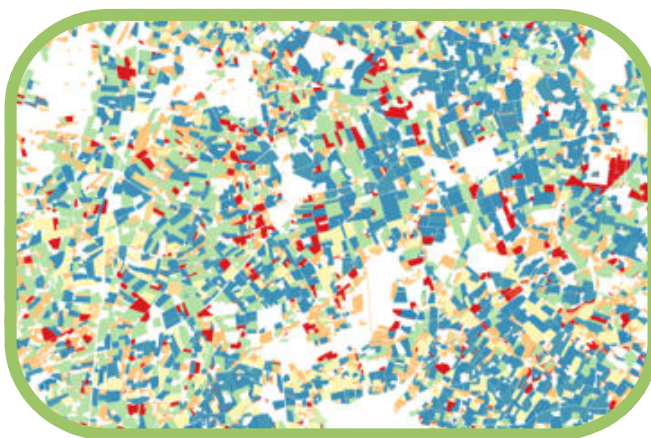
### Data

Class: Compilation of government datasets

Type: Vector

License: Creative Commons Attribution-ShareAlike 4.0 International License

DOI: <https://doi.org/10.5281/zenodo.7476474>



### Spatial

Resolution: not applicable

Extent: 17 Member States across EU (in green)

Granularity: Parcel level (sub-communal)

Coordinate System: Projected (country specific)



### Temporal

Resolution: Yearly

Extent: 2015-2023

Frequency: Annual

## DESCRIPTION

**Land Parcel Identification Systems (LPIS)** data cover the essential part of the spatial information used to support Integrated Administration and Control System (IACS) applications under the common agricultural policy (CAP). It is individually collected from the EU Member States and contains **georeferenced blocks of agricultural parcels that have been identified and are eligible for EU aid application**. This data usually shows the main crop for a certain year as the subsidy is granted concerning that.

The EuroCrops data combines all publicly available self-declared crop reporting datasets from countries of the European Union. As the raw data obtained from the countries does not come in a unified, machine-readable taxonomy, the EuroCrops data rely on a Hierarchical Crop and Agriculture Taxonomy (HCAT) that harmonises all declared crops across the European Union.





## USER COSTS

Free & Open Access for Austria, Belgium, Czech Republic, Germany, Denmark, Estonia, Spain, France, Croatia, Lithuania, Latvia, Netherlands, Portugal, Romania, Sweden, Slovenia, Slovakia.

## DATA INFRASTRUCTURE



**Infrastructure needed:** GIS (QGIS for example) or software environment for statistical computing (R for example).

**Data repositories/storage needed:** Datasets are hosted on Zenodo and made available for free.



## DATA GOVERNANCE & MANAGEMENT

**Staff costs:** Data can be easily downloaded. Its use requires a certain expertise in the world of geographic information and the structure of LPIS data (low-level GIS Technician ~25,000€ + charges per year).

**Data analysis needed:** None for simple visualisation of crops.

**Quality assurance:** Data and annotations are derived from self-declarations by farmers receiving subsidies under the Common Agricultural Policy (CAP) of the European Commission, with the usual caveat of possible errors. Moreover, this data does not represent the complete state of crops that are cultivated in each EU Member State as fields from farmers that do not receive subsidies from the CAP are not represented.

## DATA DOCUMENTATION AND/OR ASSISTANCE IN DATA USE



<https://github.com/maja601/EuroCrops>; <https://github.com/maja601/EuroCrops/wiki>;  
<https://zenodo.org/records/10118572>

### How to cite

Dataset: Schneider, M., Chan, A., & Körner, M. (2023). EuroCrops [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.10118572>



GRANULAR

DATA FICHES

# CLIMATE

BENCHMARK OF PERFORMANCE AND COSTS ENABLING  
IDENTIFICATION OF VIABLE OPTIONS GOING FORWARD







### ECMWF: ERA5 historic climate variables

<https://www.mapillary.com/>

Ian McCallum & Martin Hofer (IIASA)

#### INDICATOR CLASS: Residential



#### Data

Class: gridded

Type: API or download

License: Licence to use Copernicus Products

DOI: <https://doi.org/10.24381/cds.143582cf>



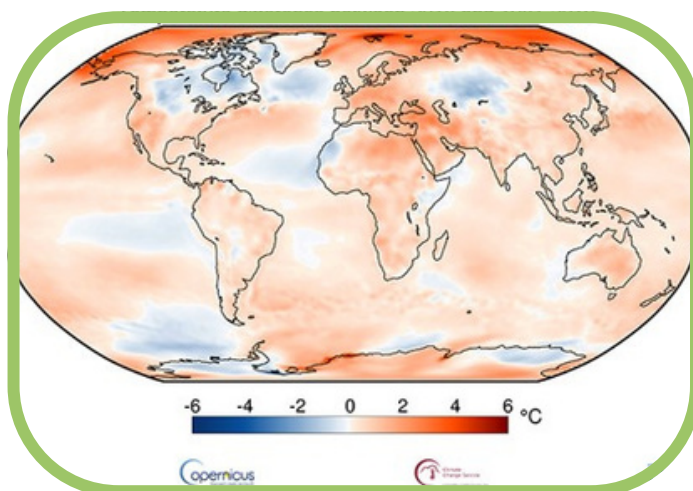
#### Spatial

Extent: Global

Granularity: 32km

Coordinate System:

Decimal Degrees, lat/lon



#### Temporal

Resolution: hourly

Extent: 1940 - today

Frequency: Daily

#### DESCRIPTION

ERA5 is the fifth generation of the European Centre for Medium-Range Weather Forecasts (ECMWF) **atmospheric reanalysis of the world's climate**. It is produced by the Copernicus Climate Change Service (C3S) at ECMWF and provides hourly estimates of a large number of **atmospheric, land and oceanic climate variables**, going from January 1940 to the present.

The data cover the Earth on a **31 km grid** and resolve the atmosphere using 137 levels from the surface up to a height of 80 km. ERA5 includes an ensemble component at half the resolution to provide information on the synoptic uncertainty of its products.



## USER COSTS

Free & open Access.

## DATA INFRASTRUCTURE



GIS software, depending on the size – more or less sophisticated data infrastructure is needed.



## DATA GOVERNANCE & MANAGEMENT

Data extraction and handling for large amounts. For smaller areas and time frames, no particular hardware is needed.

1 PM to cater data for a specific local purpose.

## DATA DOCUMENTATION AND/OR ASSISTANCE IN DATA USE



<https://confluence.ecmwf.int/display/CKB/ERA5%3A+data+documentation>

### How to cite

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Mucoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hylm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., Thépaut, J.-N. (2017): Complete ERA5 from 1940: Fifth generation of ECMWF atmospheric reanalyses of the global climate. Copernicus Climate Change Service (C3S) Data Store (CDS). DOI: 10.24381/cds.143582cf



# GRANULAR

## DATA FICHE

### Worldclim: Global climate and weather data

<https://www.worldclim.org/data/index.html>

Ian McCallum & Martin Hofer (IIASA)

#### INDICATOR CLASS: Environmental



#### Data

Class: Model-based

Type: Geo-tiff

License: Data freely available for academic and other non-commercial use.

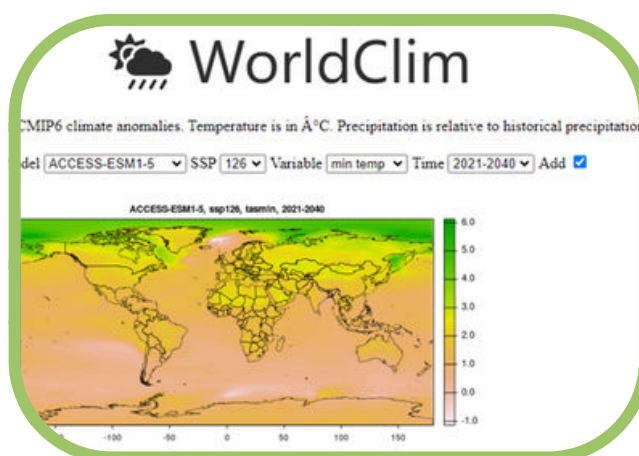


#### Temporal

Resolution: Historical climate data, Historical monthly weather data, Future climate data.

Extent: 1970-2000, 1960-2018, The monthly values were averages over 20-year periods (2021-2040, 241-2060, 2061-2080, 2081-2100).

Frequency: version 2.1 released in 2020.



#### Spatial

Resolution: 10, 5, 2.5 minutes, 30 seconds

Extent: Global

Granularity: Gridded

Coordinate System: Geographic WGS84

#### DESCRIPTION

WorldClim is a database of **high spatial resolution global weather and climate data**. This data can be used for mapping and spatial modelling.

The data is provided for use in research and related activities, and some specialised skills and knowledge are needed to use them.

Further guidance is available here: <https://www.worldclim.org/data/help.html>.



## USER COSTS

Free & Open Access.

## DATA INFRASTRUCTURE



**Infrastructure needed:** Minimal infrastructure needed. These datasets are relatively small, and can simply be downloaded directly, or can be found e.g. via R. Packages (geodata).



## DATA GOVERNANCE & MANAGEMENT

**Data analysis needed:** The files come ready to use with little reprocessing required. Nonetheless, basic GIS and climate expertise is required to fully understand and work with the files.

**Staff costs:** with basic data handling skills, this data can be acquired and processed for ingestion into models, etc. in a short time – meaning costs are negligible.

## DATA DOCUMENTATION AND/OR ASSISTANCE IN DATA USE



Good documentation is available at [worldclim.org](http://worldclim.org). Furthermore, the pre-processing effort that has gone into producing this dataset ensures that it is easy to work with.

### How to cite

Fick, S.E. and R.J. Hijmans, 2017. WorldClim 2: new 1km spatial resolution climate surfaces for global land areas. [International Journal of Climatology](#) 37 (12): 4302-4315.





GRANULAR

DATA FICHES

# DEMOGRAPHY

BENCHMARK OF PERFORMANCE AND COSTS ENABLING  
IDENTIFICATION OF VIABLE OPTIONS GOING FORWARD





GRANULAR

DATA FICHE

## GHS-POP by JRC

[https://ghsl.jrc.ec.europa.eu/ghs\\_pop2023.php](https://ghsl.jrc.ec.europa.eu/ghs_pop2023.php)

Hal Voepel (University of Southampton)

### INDICATOR CLASS: Demography, residential



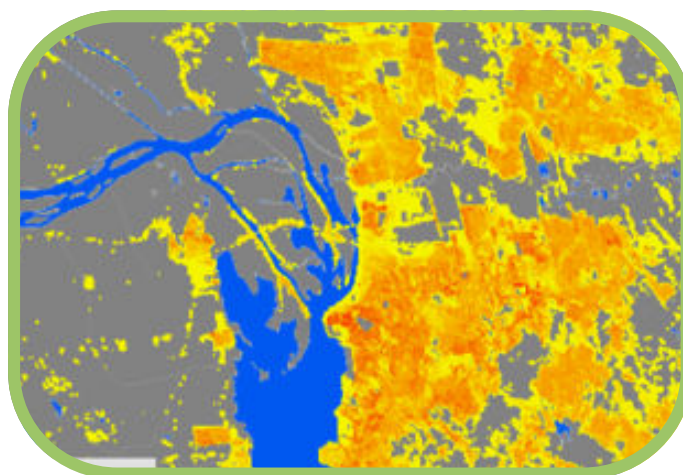
#### Data

Class: Census-based + Remote-sensing

Type: Raster, GeoTIFF

License: Full Open Access (Creative Commons Attribution 4.0 International License)

DOI:10.2905/2FF68A52-5B5B-4A22-8F40-C41DA8332CFE



#### Spatial

Resolution: 100m, 1 km, 3 arcsec, 30 arcsec

Extent: Global

Granularity: Gridded

Coordinate System: World Mollweide (ESRI: 54009) and WGS 1984 (EPSG: 4326)



#### Temporal

Resolution: Every 5 years

Extent: 1975-2030

Frequency: >1 years

### DESCRIPTION

The **spatial raster dataset** (GHS-POP – R2023A) depicts the distribution of **residential population**, expressed as the number of people per cell (float64: [0, ∞), NoData: -200).

Residential population estimates between 1975 and 2020 in 5-year intervals and projections to 2025 and 2030 derived from CIESIN GPWv4. 11 were disaggregated from census or administrative units to grid cells, informed by the distribution, volume, and classification of built-up as mapped in the Global Human Settlement Layer (GHSL) global layer per corresponding epoch. The main characteristics of this dataset are listed in the citation text.



## USER COSTS

Free & Open Access.

## DATA INFRASTRUCTURE



**Infrastructure needed:** Any raster capable GIS software.

**Data repositories/storage needed:** ~9GB for 100m global grid.



## DATA GOVERNANCE & MANAGEMENT

**Staff costs:** None.

**Data analysis needed:** None.

**Quality assurance:** See §2.5 in the report linked below.

## DATA DOCUMENTATION AND/OR ASSISTANCE IN DATA USE



The complete information about the GHSL main products can be found in the [GHSL Data Package 2023](#) report.

### How to cite

Dataset:

Schiavina M., Freire S., Carioli A., MacManus K. (2023):

GHS-POP R2023A - GHS population grid multitemporal (1975-2030). European Commission, Joint Research Centre (JRC)

PID: [Dataset](#)

doi: <https://doi.org/10.2905/2FF68A52-5B5B-4A22-8F40-C41DA8332CFE>

Concept & Methods:

Freire S., MacManus K., Pesaresi M., Doxsey-Whitfield E., Mills J. (2016)

Development of new open and free multi-temporal global population grids at 250 m resolution. Geospatial Data in a Changing World; Association of Geographic Information Laboratories in Europe (AGILE), AGILE 2016



GRANULAR

## DATA FICHE

# WorldPop – Population Weighted Density by the University of Southampton

<https://hub.worldpop.org/project/categories?id=20>

Hal Voepel (University of Southampton)

### INDICATOR CLASS: Demography, residential



#### Data

Class: Census-based + Remote-sensing and GIS reference data

Type: Raster

License: Full Open Access

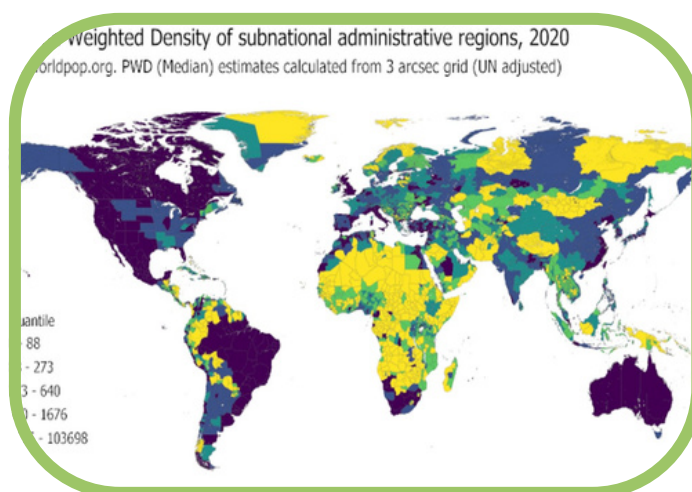
DOI by resolution and coverage:

[10.5258/SOTON/WP00701](https://doi.org/10.5258/SOTON/WP00701) (100m national)

[10.5258/SOTON/WP00702](https://doi.org/10.5258/SOTON/WP00702) (1km national)

[10.5258/SOTON/WP00703](https://doi.org/10.5258/SOTON/WP00703) (100m subnational)

[10.5258/SOTON/WP00704](https://doi.org/10.5258/SOTON/WP00704) (1km subnational)



#### Spatial

Resolution: 100m, 1 km

Extent: Global

Granularity: Gridded

Coordinate System: World Mollweide (ESRI: 54009) and WGS 1984 (EPSG: 4326)



#### Temporal

Resolution: Every 5 years

Extent: 2005-2020

Frequency: >1 years

### DESCRIPTION

**Population Weighted Density (PWD)** is an alternative to conventional approaches to population density that is arguably better suited to some types of research in the fields of social science and epidemiology. In this release, WorldPop publishes what we believe may be the first set of **global estimates for PWD**, which has been offered at national and subnational levels since 2000. Data are available at both national and subnational levels.

Please make sure you read the Population Weighted Density overview page before choosing and downloading a PWD dataset (see link to Methods at the end of this document).





## USER COSTS

Free & Open Access.

## DATA INFRASTRUCTURE



**Infrastructure needed:** any raster capable GIS.

**Data repositories/storage needed:** < 1MB.



## DATA GOVERNANCE & MANAGEMENT

**Staff costs:** None.

**Data analysis needed:** None.

**Quality assurance:** See cited references and methods below.

## DATA DOCUMENTATION AND/OR ASSISTANCE IN DATA USE



Methods: <https://www.worldpop.org/methods/pwd/>

### How to cite

Dataset: Robin Edwards, Maksym Bondarenko, Andrew J. Tatem and Alessandro Sorichetta. Unconstrained subnational Population Weighted Density in 2000, 2005, 2010, 2015 and 2020 ( 100m resolution ). WorldPop, University of Southampton, UK. doi:10.5258/SOTON/WP00703

Concept & Methods:

Freire, Sergio; MacManus, Kytt; Pesaresi, Martino; Doxsey-Whitfield, Erin; Mills, Jane (2016): Development of new open and free multi-temporal global population grids at 250 m resolution. Geospatial Data in a Changing World; Association of Geographic Information Laboratories in Europe (AGILE). AGILE 2016.



# GRANULAR DATA FICHE

## WorldPop – Population Density by University of Southampton

<https://hub.worldpop.org/project/categories?id=18>

Hal Voepel (University of Southampton)

### INDICATOR CLASS: Demography, residential



#### Data

Class: Census-based + Remote-sensing and GIS reference data

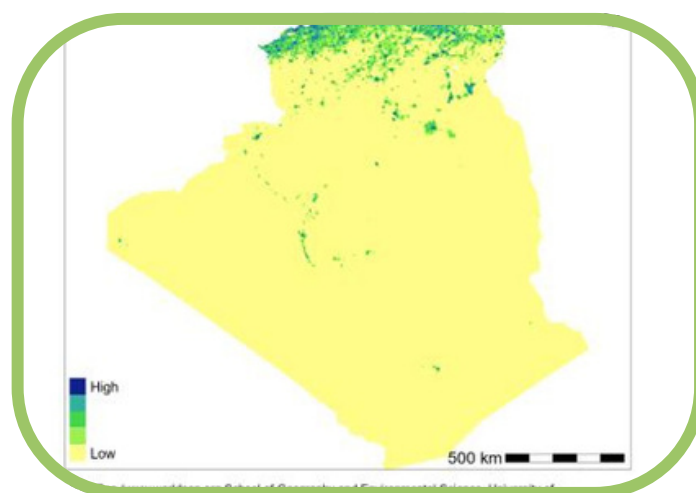
Type: Raster

License: Full Open Access

DOI by resolution and method:

[10.5258/SOTON/WP00674](https://doi.org/10.5258/SOTON/WP00674) (1km individual countries)

[10.5258/SOTON/WP00675](https://doi.org/10.5258/SOTON/WP00675) (1km individual countries, UN adjusted)



#### Spatial

Resolution: 1 km

Extent: Country-level

Granularity: Gridded

Coordinate System: WGS 1984 (EPSG: 4326)



#### Temporal

Resolution: Annually

Extent: 2000-2020

Frequency: >1 years

### DESCRIPTION

**Population Weighted Density (PWD)** is an alternative to conventional approaches to population density that is arguably better suited to some types of research in the fields of social science and epidemiology. In this release, WorldPop publishes what we believe may be the first set of global estimates for PWD, which has been offered at national and subnational levels since 2000. Data are available at both national and subnational levels.

Please make sure you read the Population Weighted Density overview page before choosing and downloading a PWD dataset (see link to Methods at the end of this document).



## USER COSTS

Free & Open Access.

## DATA INFRASTRUCTURE



**Infrastructure needed:** any raster capable GIS.

**Data repositories/storage needed:** < 1MB.



## DATA GOVERNANCE & MANAGEMENT

**Staff costs:** None.

**Data analysis needed:** None.

**Quality assurance:** See cited references and methods below.

## DATA DOCUMENTATION AND/OR ASSISTANCE IN DATA USE



Methods: <https://www.worldpop.org/methods/pwd/>

### How to cite

Dataset: WorldPop ([www.worldpop.org](http://www.worldpop.org) - School of Geography and Environmental Science, University of Southampton; Department of Geography and Geosciences, University of Louisville; Departement de Geographie, Universite de Namur) and Center for International Earth Science Information Network (CIESIN), Columbia University (2018). Global High Resolution Population Denominators Project - Funded by The Bill and Melinda Gates Foundation (OPP1134076).

Concept & Methods: Sorichetta, A., Hornby, G., Stevens, F. et al. High-resolution gridded population datasets for Latin America and the Caribbean in 2010, 2015, and 2020. Sci Data 2, 150045 (2015). <https://doi.org/10.1038/sdata.2015.45>



GRANULAR

## DATA FICHE

# WorldPop – Age & Sex Structures by University of Southampton

<https://hub.worldpop.org/project/categories?id=8>

Hal Voepel (University of Southampton )

## INDICATOR CLASS: Demography, residential



### Data

Class: Census-based + Remote-sensing and GIS reference data

Type: Raster

License: Open Access (Creative Commons Attribution 4.0 International License)

DOI by group, method, and resolution:

[10.5258/SOTON/WP00696](#) Constrained countries in 2020 (100m)

[10.5258/SOTON/WP00699](#) Constrained countries in 2020 UN adjusted (100m)

[10.5258/SOTON/WP00732](#) School age population (1km)

[10.5258/SOTON/WP00654](#) Unconstrained global mosaic 2000-2020 (1km)

[10.5258/SOTON/WP00646](#) Unconstrained single countries 2000-2020 (100m)

No DOI assigned to data: Women of childbearing age (15-49yrs) in 2015 (1km)



### Spatial

Resolution: 100m, 1 km

Extent: Global and country-level

Granularity: Gridded

Coordinate System: WGS 1984 (EPSG: 4326)



### Temporal

Resolution: Every 5 years

Extent: 2000-2020 (unconstrained), 2020 (constrained and school age), and 2015 (women of childbearing age)

Frequency: >1 years

## DESCRIPTION

WorldPop produces a variety of **gridded age and sex structures**. A description of the modelling methods used for age and sex structures can be found in Tatem et.al. (2013) and Pezzulo et.al. (2017). Both top-down constrained and unconstrained versions of the dataset are available, and the difference between the two methods are outlined [here](https://www.worldpop.org/methods/top_down_constrained_vs_unconstrained/): ([https://www.worldpop.org/methods/top\\_down\\_constrained\\_vs\\_unconstrained/](https://www.worldpop.org/methods/top_down_constrained_vs_unconstrained/)).

The datasets represent the outputs from a project focused on the construction of consistent 100m resolution population count datasets for all countries by male/female in 5-year classes with a final class for infants defined as less than one year old.





## USER COSTS

Free & Open Access.

## DATA INFRASTRUCTURE



**Infrastructure needed:** Any raster capable GIS.

**Data repositories/storage needed:** Up to 3+GB global tile.



## DATA GOVERNANCE & MANAGEMENT

**Staff costs:** None.

**Data analysis needed:** None.

**Quality assurance:** See population methods below.

## DATA DOCUMENTATION AND/OR ASSISTANCE IN DATA USE



**Methods:** <https://hub.worldpop.org/project/categories?id=3>

### How to cite

Concept & Methods:

Pezzulo, C., Hornby, G., Sorichetta, A. et al. Sub-national mapping of population pyramids and dependency ratios in Africa and Asia. *Sci Data* 4, 170089 (2017). <https://doi.org/10.1038/sdata.2017.89>.

Tatem, A.J., Garcia, A.J., Snow, R.W. et al. Millennium development health metrics: where do Africa's children and women of childbearing age live?. *Popul Health Metrics* 11, 11 (2013). <https://doi.org/10.1186/1478-7954-11-11>.



GRANULAR

## DATA FICHE

# WorldPop – Population Counts by University of Southampton

<https://hub.worldpop.org/project/categories?id=3>

Hal Voepel (University of Southampton )

### INDICATOR CLASS: Demography, residential



#### Data

Class: Census-based + Remote-sensing and GIS reference data

Type: Raster

License: Full Open Access (Creative Commons Attribution 4.0)

DOI by resolution and method:

[10.5258/SOTON/WP00682](https://doi.org/10.5258/SOTON/WP00682) (100m constrained countries)

[10.5258/SOTON/WP00683](https://doi.org/10.5258/SOTON/WP00683) (100m constrained countries, UN adjusted)

[10.5258/SOTON/WP00647](https://doi.org/10.5258/SOTON/WP00647) (1km unconstrained global mosaic)

[10.5258/SOTON/WP00645](https://doi.org/10.5258/SOTON/WP00645) (100m unconstrained countries)

[10.5258/SOTON/WP00670](https://doi.org/10.5258/SOTON/WP00670) (1km unconstrained countries)

[10.5258/SOTON/WP00660](https://doi.org/10.5258/SOTON/WP00660) (100m unconstrained countries, UN adjusted)

[10.5258/SOTON/WP00671](https://doi.org/10.5258/SOTON/WP00671) (1km unconstrained countries, UN adjusted)

[wopr.worldpop.org/](https://wopr.worldpop.org/) (Bespoke methods for individual countries)



#### Temporal

Resolution: Annually

Extent: 2000-2020 (Unconstrained) 2020 (Constrained)

Frequency: >1 years



#### Spatial

Resolution: 1 km

Extent: Country-level

Granularity: Gridded

Coordinate System: WGS 1984 (EPSG: 4326)

### DESCRIPTION

Bespoke methods used to produce datasets for specific individual countries are available through the **WorldPop Open Population Repository** (WOPR) link above under Data. These are 100m resolution gridded population estimates using the bottom-up or top-down methods described here: <https://www.worldpop.org/methods/populations/>. The remaining datasets linked above use the top-down method, with either constrained or unconstrained top-down aggregation methods. Be sure to read the [Top-Down estimate modelling overview page](#) to see which dataset is appropriate for your purposes. Datasets are available as GeoTIFF or ASCII XYZ at 100m and 1km resolution.



## USER COSTS

Free & Open Access.

## DATA INFRASTRUCTURE



**Infrastructure needed:** Any raster capable GIS.

**Data repositories/storage needed:** < 1MB.



## DATA GOVERNANCE & MANAGEMENT

**Staff costs:** None.

**Data analysis needed:** None.

**Quality assurance:** See cited references and methods below.

## DATA DOCUMENTATION AND/OR ASSISTANCE IN DATA USE



**Methods:** <https://hub.worldpop.org/project/categories?id=3>

### How to cite

Dataset: : Bondarenko M., Kerr D., Sorichetta A., and Tatem, A.J. 2020. Census/projection-disaggregated gridded population datasets for 51 countries across sub-Saharan Africa in 2020 using building footprints. WorldPop, University of Southampton, UK. doi:10.5258/SOTON/WP00682 (see citations for each individual link in blue Data section above)

Concept & Methods: Stevens FR, Gaughan AE, Linard C, Tatem AJ (2015) Disaggregating Census Data for Population Mapping Using Random Forests with Remotely-Sensed and Ancillary Data. PLoS ONE 10(2): e0107042. <https://doi.org/10.1371/journal.pone.0107042>



GRANULAR

## DATA FICHE

# Data for Good: HiRes Population Density

<https://dataforgood.facebook.com/>

Hal Voepel (University of Southampton)

### INDICATOR CLASS: Demography, residential



#### Data

Class: Census-based + Remote-sensing

Type: Raster, GeoTIFF

License: Full Open Access (Creative Commons Attribution 4.0)



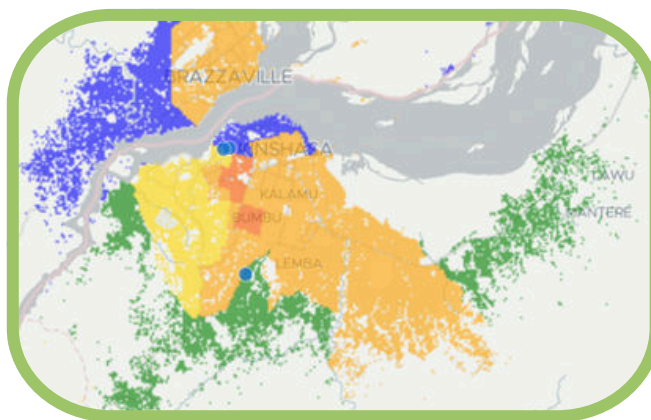
#### Spatial

Resolution: 30m

Extent: Global

Granularity: Gridded by country

Coordinate System: WGS 1984 (EPSG: 4326)



#### Temporal

Resolution: Varies

Extent: Varies

Frequency: ~ Annually

### DESCRIPTION

The **DataforGood high-resolution maps** estimate the number of people living within 30-meter grid tiles in nearly every country around the world. These datasets also provide insights into the **distribution of certain populations within each country**, including the number of children under five, the number of women of reproductive age, as well as young and elderly population, at unprecedentedly high resolutions.

As an example, these maps can provide the necessary insights for **health organisations** to allocate resources and control outbreaks. There are many use cases for understanding the demographics of various populations – demographics can help organisations target vaccination campaigns, plan infrastructure, and distribute resources.

These maps are not built using Facebook data and instead rely on **combining the power of machine vision AI with satellite imagery and census information**.





## USER COSTS

Free & Open Access.

## DATA INFRASTRUCTURE



**Infrastructure needed:** Any raster-capable GIS software.

**Data repositories / storage needed:** <100MB for 30m UK data.



## DATA GOVERNANCE & MANAGEMENT

**Staff costs:** None.

**Data analysis needed:** None.

**Quality assurance:** See individual dataset documentation.

## DATA DOCUMENTATION AND/OR ASSISTANCE IN DATA USE



See documentation for each dataset here: [https://data.humdata.org/organization/meta?q=population%20density&sort=if\(gt\(last\\_modified%2Creview\\_date\)%2Clast\\_modified%2Creview\\_date](https://data.humdata.org/organization/meta?q=population%20density&sort=if(gt(last_modified%2Creview_date)%2Clast_modified%2Creview_date)

### How to cite

Dataset: Varies by location.

Concept & Methods: Tiecke, T.G., Liu, X., Zhang, A., Gros, A., Li, N., Yetman, G., Kilic, T., Murray, S., Blankespoor, B., Prydz, E.B., & Dang, H.H. (2017). Mapping the world population one building at a time. [ArXiv:1712.05839](https://arxiv.org/abs/1712.05839).



# GRANULAR

## DATA FICHE

### ARDECO Population Series

<https://urban.jrc.ec.europa.eu/ardeco?lng=en>

Hal Voepel (University of Southampton)

#### INDICATOR CLASS: Demography, administrative level



#### Data

Class: Census-based + growth model  
Type: CSV with NUTS/LAU primary key  
License: Full Open Access



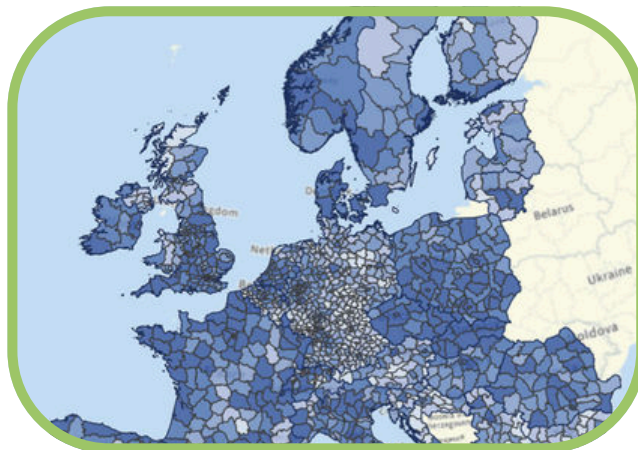
#### Temporal

Resolution: Annual  
Extent: 1960 to current + 2 years forecast  
Frequency: updated biannually



#### Spatial

Resolution: NUTS / LAU  
Extent: EU27 + others, 50 countries  
Granularity: NUTS/LAU  
Coordinate System: None



#### DESCRIPTION

The “**Annual Regional Database of the European Commission**” (ARDECO) is maintained by the Joint Research Centre, in close collaboration with the Directorate General for Regional and Urban Policy. Data contained in ARDECO are a set of long-time series based on the ESA 2010 system for the last period and on ESA 95 and ESA 79 for the earlier years, where growth rates are used to gap-fill missing values.

The database covers - at an annual scale -, a total of **50 countries**, of which 27 EU countries and 23 non-EU countries. Data are in tabular form with NUTS/LAU primary keys. ARDECO is **updated biannually** with major updates released in early March (after the release of updated regional accounts by Eurostat) and in May and November, following the release of Directorate General ECFIN's short-term economic forecasts. ARDECO data are **publicly available** and **free of charge**.



## USER COSTS

Free of charge & Access limited (request form).

## DATA INFRASTRUCTURE



**Infrastructure needed:** Excel or GIS software.

**Data repositories/storage needed:** 1.5MB.



## DATA GOVERNANCE & MANAGEMENT

**Staff costs:** None.

**Data analysis needed:** None.

**Quality assurance:** see documentation [here](#).

## DATA DOCUMENTATION AND/OR ASSISTANCE IN DATA USE



See documentation for each dataset [here](#).

### How to cite:

IDataset: Metadata. [https://economy-finance.ec.europa.eu/system/files/2022-10/Reference%20Metadata%20AMECO\\_September%202022.pdf](https://economy-finance.ec.europa.eu/system/files/2022-10/Reference%20Metadata%20AMECO_September%202022.pdf)



GRANULAR

DATA FICHES

# DIGITALISATION

BENCHMARK OF PERFORMANCE AND COSTS ENABLING  
IDENTIFICATION OF VIABLE OPTIONS GOING FORWARD





GRANULAR

## DATA FICHE

# Rural household internet access by Eurostat

[https://ec.europa.eu/eurostat/databrowser/view/isoc\\_ci\\_in\\_h\\_\\_custom\\_8667150/default/table?lang=enL:](https://ec.europa.eu/eurostat/databrowser/view/isoc_ci_in_h__custom_8667150/default/table?lang=enL)

Lauri Niskanen (LUKE)

## INDICATOR CLASS: Residential, Digitalisation



### Data

Class: Statistics

Type: Tabular

License: <https://ec.europa.eu/eurostat/web/main/about-us/policies/copyright>.



### Spatial

Extent: EU

Coordinate System: Projected (country specific)



### Temporal

Resolution: Yearly

Extent: 2002 - 2022

Frequency: Annual

## DESCRIPTION

The **EU survey on the use of Information and Communication Technologies (ICT) in households and by individuals** is an annual survey conducted since 2002. It aims at collecting and disseminating harmonised and comparable information on the use of ICT in households and by individuals.

Data presented in this domain is collected every year by the National Statistical Institutes and is based on Eurostat's annual model questionnaire. This questionnaire is updated each year to reflect the evolving situation of information and communication technologies.





## USER COSTS

Free & Open Access.

## DATA INFRASTRUCTURE



**Infrastructure needed:** Software environment of statistics in XLS, CVS, TSV and other format.

**Data repositories:** FTPStore.



## DATA GOVERNANCE & MANAGEMENT

**Staff costs:** Basic statistical skills or more advanced, depending on usage.

**Data analysis needed:** Depending on the analysis. Charts and graphs can be viewed directly.

**Quality assurance:** The content of metadata and quality reports follows the Commission Implementing Regulation (EU) 2019/2180 of 16 December 2019 specifying the detailed arrangements and content for the quality reports under Regulation (EU) 2019/1700 of the European Parliament and of the Council.

## DATA DOCUMENTATION AND/OR ASSISTANCE IN DATA USE



[https://ec.europa.eu/eurostat/databrowser/view/isoc\\_ci\\_in\\_h\\_\\_custom\\_8667150/default/table?lang=en](https://ec.europa.eu/eurostat/databrowser/view/isoc_ci_in_h__custom_8667150/default/table?lang=en)  
[https://ec.europa.eu/eurostat/cache/metadata/en/isoc\\_i\\_esms.htm](https://ec.europa.eu/eurostat/cache/metadata/en/isoc_i_esms.htm)

### How to cite

Dataset: Eurostat. ICT usage in households and by individuals (isoc\_i): reference metadata in Euro SDMX Metadata Structure (ESMS).

"This study/report/paper is based on data from Eurostat, EU Statistics Rural household internet access in [reference year(s)]."  
Concept & Methodology: [https://ec.europa.eu/eurostat/cache/metadata/en/isoc\\_i\\_esms.htm](https://ec.europa.eu/eurostat/cache/metadata/en/isoc_i_esms.htm)



GRANULAR

DATA FICHES

# ECONOMIC DEVELOPMENT

BENCHMARK OF PERFORMANCE AND COSTS ENABLING  
IDENTIFICATION OF VIABLE OPTIONS GOING FORWARD





GRANULAR

DATA FICHE

# Tripadvisor: Hotel, Restaurant, Attractions and visitors statistics

<https://www.tripadvisor.com/developers>

Ian McCallum & Martin Hofer (IIASA)

## INDICATOR CLASS: Residential



### Data

Class: GIS & tabular

Type: API

License: Proprietary API

Use Licence <https://tripadvisor-content-api.readme.io/reference/terms-of-use#3-licensed-content>



### Spatial

Extent: Global

Granularity: street-level

CoordinateSys.: WGS 84

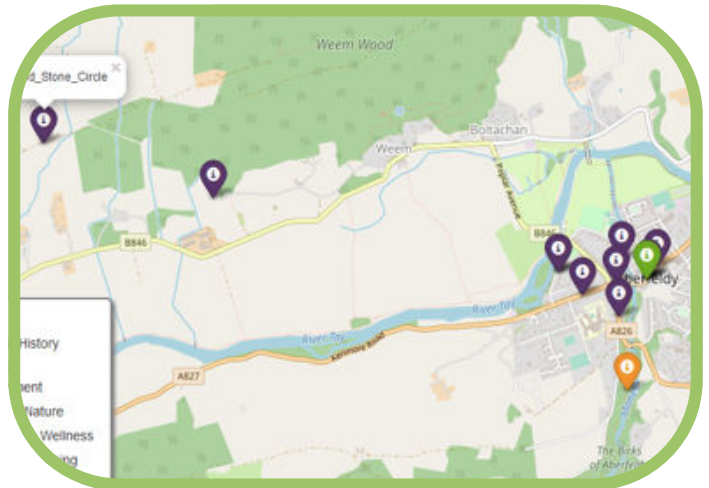


### Temporal

Resolution: Monthly

Extent: 2015 -

Frequency: realtime



## DESCRIPTION

**TripAdvisor** is a popular travel website that offers **user-generated reviews and comparisons** for hotels, restaurants, and attractions worldwide.

Launched in 2000, it has become a go-to platform for travellers to plan and book their trips, as well as share experiences, helping others make informed travel decisions. It is one of the largest travel communities globally, impacting travel industry dynamics significantly.



## USER COSTS

Proprietary API.

## DATA INFRASTRUCTURE



Database for all attractions and reviews for Europe would be around 20 GB. A relation database (SQL) with locations, reviews users and tags should suffice.



## DATA GOVERNANCE & MANAGEMENT

The tags for the locations need to be consolidated into rural functional areas. Manually or through the use of large language models.

Getting location data of activities/hotels via the API 1PM.

With additional analysis the reviews can inform tourist statistics. The data needs to be calibrated against other tourism statistics first. (2PM)

## DATA DOCUMENTATION AND/OR ASSISTANCE IN DATA USE



<https://tripadvisor-content-api.readme.io>

### How to cite

<https://tripadvisor-content-api.readme.io>



GRANULAR

## DATA FICHE

# Airbnb/Hotel locations and capacity

<http://insideairbnb.com/data-requests>

Ian McCallum & Martin Hofer (IIASA)

### INDICATOR CLASS: Residential



#### Data

Class: GIS points & tabular

Type: table



#### Spatial

Extent: Global

Granularity: street-level

Coordinate System: WGS 84



#### Temporal

Resolution: 30, 60, 90, 365 days

Extent: 2015

Frequency: annual



### DESCRIPTION

As reported on Inside Airbnb's website: "Inside Airbnb is a **mission-driven project** that provides **data and advocacy about Airbnb's impact on residential communities.**"

Inside Airbnb scrapes Airbnb data and gives access to the public. They provide the location (longitude/latitude), room and accommodation type, information on the accommodation, and the number of available rooms in the next 30, 60, 90 and 365 days, from the last scraping.

Access to the data varies. Many cities are available for free. Archived data or new locations require a budget, depending on how much the requesters' mission aligns with Inside Airbnb's mission.





## USER COSTS

Sold or shared on request.

## DATA INFRASTRUCTURE



A GIS-capable database for all of Europe would be fairly easily manageable (e.g., Postgres with a GIS plugin).



## DATA GOVERNANCE & MANAGEMENT

Data comes very well-cleaned and prepared. Not much work is needed. Base knowledge is needed.

## DATA DOCUMENTATION AND/OR ASSISTANCE IN DATA USE



<https://docs.google.com/spreadsheets/d/1iWCNJcSutYqpULSQHINyGlnUvHg2BoUGoNRIGa6Szc4/htmlview#gid=1322284596>





# GRANULAR

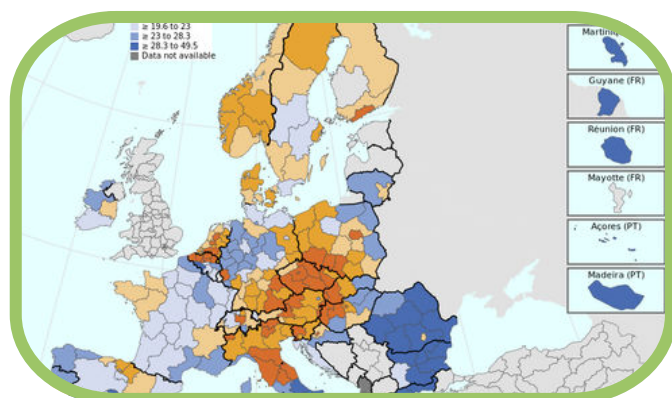
## DATA FICHE

### EU-SILC microdata

<https://ec.europa.eu/eurostat/web/microdata/european-union-statistics-on-income-and-living-conditions>

Carlos Tapia (Nordregio)

#### INDICATOR CLASS: Productive



People at risk of poverty or social exclusion by NUTS 2 regions



#### Data

Class: Alphanumeric

Type: Tabular

License: EU-SILC data belongs to the National Statistical Institutes (NSI) of EU Member States (MS). Licensing varies between countries. Eurostat performs a first evaluation of each research proposal and then consults with NSI before the data is made accessible to researchers. Datasets contain partially anonymised microdata delivered as scientific use files that can be downloaded and used by authorised researchers under restrictive Terms of Use. There is also a public version of microdata files.



#### Spatial

Extent: EU (27 countries) + UK, CH, IS, NO, RS.

Granularity: IESS Regulation (EU:2019/1700) states that reliable statistics relating to persons and households based on individual samples should be provided at the NUTS-2 level. However, the availability of data at the NUTS-2 level is limited by effective sample sizes. The estimated ratio of people at risk of poverty or social exclusion to the population is not compulsory for NUTS-2 regions with fewer than 0.5 million inhabitants. nor for NUTS-1 regions with under 100,000 inhabitants. Even in the most recent EU-SILC surveys, MS continue to report at-risk-of-poverty or social exclusion data at the NUTS-1 level, in some cases, only at national level.



#### Temporal

Resolution: Annual, 5-annual (seenote)

Extent: 2004-2022

Frequency: Yearly

#### DESCRIPTION

The **EU Statistics on Income, Social Inclusion and Living Conditions (EU-SILC)** is a cross-sectional and longitudinal sample survey, coordinated by Eurostat, based on data from the EU Member States. The EU-SILC survey covers objective and subjective aspects of these themes in both **monetary and non-monetary terms for households and individuals**.

The EU-SILC provides two types of annual survey data:

- Cross-sectional data with variables on income, poverty, social exclusion and other living conditions
- Longitudinal data about individual-level changes over time, observed periodically over 4 years.



## USER COSTS

Free & Open Access.

## DATA INFRASTRUCTURE



The processing of Eurostat microdata does not require a very sophisticated or particularly powerful computing capacity. However, the strict security and confidentiality requirements set by Eurostat may increase data management costs. To fulfil the Terms of Use set by Eurostat, the following requirements need to be fulfilled:

- Confidential data for scientific purposes must be stored on a password-protected computer.
- Access to the data must be restricted to authorised researchers named in the research proposal.
- Intermediate results of analysis containing confidential data must be stored in a protected environment.
- The confidential data for scientific purposes must be solely used on the premises of the research entity.

These terms prevent a distributed use of the microdata, and also encourage the establishment of a secured IT environment for microdata manipulation, with devoted working stations with physical access restrictions.



## DATA GOVERNANCE & MANAGEMENT

No specific staff costs related to the use of EU-SILC microdata files. Still, the EU-SILC SUFs provide access to complex data survey. Adequate use of these files is only possible for researchers with a robust knowledge of statistics and data management capacity.

## DATA DOCUMENTATION AND/OR ASSISTANCE IN DATA USE



Upon download of the EU-SILC data, the users are provided with several documents describing the SUF files.

Several specialised forums and conferences gather social researchers who use EU-SILC in research. Additionally, the EU Platform "Collaboration in Research and Methodology for Official Statistics" curates a Forum where users can pose questions related to EU-SILC scientific use files: [https://cros-legacy.ec.europa.eu/EU-SILC-SUF/forum\\_en](https://cros-legacy.ec.europa.eu/EU-SILC-SUF/forum_en).

The EU-SILC user conference: <https://www.statistik.at/en/about-us/events/eu-silc-user-conference-2023>

### How to cite

Researchers must cite the source of EU-SILC data as follows: "This study/report/paper is based on data from Eurostat, EU Statistics on Income and Living Conditions [reference year(s)]." The following disclaimer must be added: "The responsibility for all conclusions drawn from the data lies entirely with the author(s)."



GRANULAR

## DATA FICHES

# ENERGY

BENCHMARK OF PERFORMANCE AND COSTS ENABLING  
IDENTIFICATION OF VIABLE OPTIONS GOING FORWARD







GRANULAR

## DATA FICHE

# Suitability map for solar energy (PV)s by JRC

URL: <https://energy-industry-geolab.jrc.ec.europa.eu/>

Lauri Niskanen (LUKE)

## INDICATOR CLASS: Energy, Productive



### Data

Class: Heterogenous

Type: Raster

License: The data in the Energy & Industry Geography Lab are free for public use

DOI: <https://doi.org/10.1016/j.enpol.2015.10.004>



### Spatial

Resolution: 1 km

Extent: EU

Coordinate System: Spherical/Web Mercator

EPSG:3857



### Temporal

Resolution: Every five years

Extent: 2010-2050

Frequency: Irregular

## DESCRIPTION

**SOLAR** (PV and CSP) data set of ENSPRESO - an open data, EU-28 wide, transparent and coherent database of **wind, solar and biomass energy potential**. Solar radiation can be converted into sustainable-produced electricity by using photovoltaic (PV) technology.

**Large-scale photovoltaic** (PV) systems provide significant environmental benefits and advantages when compared to conventional, non-renewable energy sources, with the reduction of greenhouse gas emissions, and the reuse of marginal lands being two key examples (IPCC, 2011).

- A European suitability map for the solar energy (PV) systems deployment is created.
- PV systems can contribute to sustainable energy production in many regions in the EU.
- There is no correlation between the EU investment and the suitability of solar energy.
- Using marginal lands to place PV systems might avoid the uptake of agricultural lands.
- Validation of the EU suitability map demonstrated a satisfactory degree of accuracy.





## USER COSTS

Free & Open Access.

## DATA INFRASTRUCTURE



**Infrastructure needed:** GIS, statistical or remote sensing software.

**Data repositories/storage needed:** Yes.



## DATA GOVERNANCE & MANAGEMENT

**Staff costs:** Basic GIS or statistical skills, more advanced depending on usage.

**Data analysis needed:** No.

**Quality assurance:** None specified.

## DATA DOCUMENTATION AND/OR ASSISTANCE IN DATA USE



<https://data.jrc.ec.europa.eu/dataset/18eb348b-1420-46b6-978a-fe0b79e30ad3#dataaccess>

### How to cite

Dataset: European Commission, Joint Research Centre (JRC) (2019): ENSPRESO - SOLAR - PV and CSP. European Commission, Joint Research Centre (JRC) [Dataset] PID: <http://data.europa.eu/89h/18eb348b-1420-46b6-978a-fe0b79e30ad3>  
Concept & Methodology: Castillo, C. P., e Silva, F. B., & Lavalle, C. (2016). An assessment of the regional potential for solar power generation in EU-28. Energy policy, 88, 86-99. <https://doi.org/10.1016/j.enpol.2015.10.004>



GRANULAR

## DATA FICHES

# HEALTH

BENCHMARK OF PERFORMANCE AND COSTS ENABLING  
IDENTIFICATION OF VIABLE OPTIONS GOING FORWARD





GRANULAR

## DATA FICHE

# Healthcare services location & number of beds in Europe by Eurostat

<https://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/healthcare-services>

Antti Hiltunen (LUKE)

## INDICATOR CLASS: Health, residential



### Data

Class: Point locations, In-situ

Type: Vector

License: Open Access (Eurostat's general copyright notice and license)



### Spatial

Resolution: Point locations

Extent: EU + Switzerland/Norway

Granularity: National

Coordinate System: WGS 1984 (EPSG: 4326)



### Temporal

Resolution: Annual

Extent: 2017-2022

Frequency: >1 years

## DESCRIPTION

The dataset contains information on the main **healthcare services** considered to be 'hospitals' by EU Member States. The definition varies slightly from country to country but roughly includes the following: "Hospitals' comprises licensed establishments primarily engaged in providing medical, diagnostic, and treatment services that include physician, nursing, and other health services to in-patients and the specialised accommodation services required by inpatients. Hospitals may also provide outpatient services as a secondary activity. Hospitals provide in-patient health services, many of which can only be provided using the specialised facilities and equipment that form a significant and integral part of the production process. In some countries, health facilities need in addition a minimum size (such as the number of beds) to be registered as a hospital." In some cases, facilities without in-patient services may be included. **A variety of data sources are available at the Member State level with different quality criteria:** level of detail, timeliness, update frequency, etc. This methodology aims to progressively build pan-EU datasets from Member State official data and sustainably maintain them. Different data sources could be selected depending on the needs and the cases but priority is given to data sourced from Health Ministries, or other official outlets indicated by Ministries. Data is semantically harmonised with a common (simple) schema. The production relies on automated processes: Member State data are retrieved, combined, harmonised and updated automatically, as far as possible, using up-to-date techniques.



## USER COSTS

Free & Open Access.

## DATA INFRASTRUCTURE



**Infrastructure needed:** GIS statistical software.

**Data repositories/storage needed:** No, .json API available.



## DATA GOVERNANCE & MANAGEMENT

**Staff costs:** Basic GIS skills.

**Data analysis needed:** No.

**Quality assurance:** Data quality varies from country to country.

## DATA DOCUMENTATION AND/OR ASSISTANCE IN DATA USE



<https://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/healthcare-services>

<https://gisco-services.ec.europa.eu/pub/healthcare/metadata.pdf>

### How to cite

Dataset: Eurostat. (2023). Healthcare services locations. Retrieved from Eurostat website:

<https://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/healthcare-services>

Concept & Methodology: Eurostat. (2020). Healthcare services in Europe. Retrieved from Eurostat website:

<https://gisco-services.ec.europa.eu/pub/healthcare/metadata.pdf>





GRANULAR

DATA FICHES

# INFRASTRUCTURE

BENCHMARK OF PERFORMANCE AND COSTS ENABLING  
IDENTIFICATION OF VIABLE OPTIONS GOING FORWARD







# GRANULAR

## DATA FICHE

# OpenStreetMap for Points of Interest

<https://www.openstreetmap.org/>

Ronan Ysebaert (CNRS)

## INDICATOR CLASS: Infrastructures



### Data

Class: pointlocations

Type: vector (nodes, ways, relations)

License: Data Commons Open Database license



### Spatial

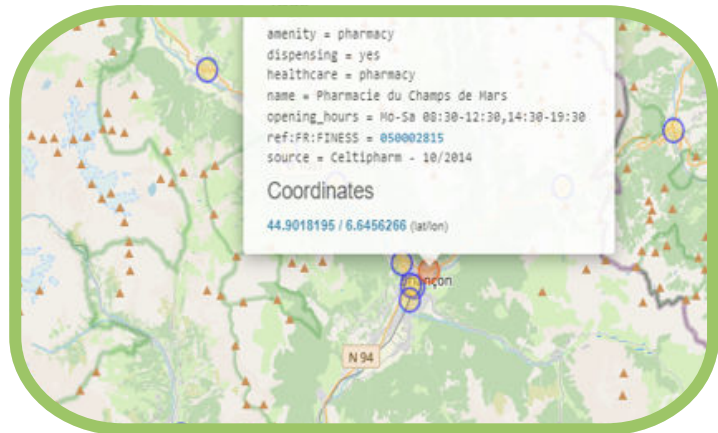
Resolution: Not relevant

Extent: World

Granularity: Not relevant

Coordinate System:

WGS 1984 (EPSG: 4326)



### Temporal

Resolution: Up-to-date

Extent: Not relevant

Frequency: Continuously updated

## DESCRIPTION

OpenStreetMap (OSM) is a well-known **crowd-sourced project** which aims to produce a **free vector geographic database** of the world. The OSM database is one of the pioneering and best-known sources of Volunteered Geographic Information (VGI). Free and open source since its creation in 2004, a **variety of data** can be found in OSM including buildings, land use, road network. Many use cases can be made from OSM (map OSM tiles, routing, etc.). Extracting Points of Interest derived from this database is one of them.

OSM represents physical features on the ground (e.g., roads or buildings) using tags attached to its basic data structure. Each tag defined by a key-value combination describes a geographic attribute of the feature being shown by that specific node, way or relation. OSM **free tagging system** allows the map to include an unlimited number of attributes describing each feature. The community agrees on certain key and value combinations for the most commonly used tags, which act as informal standards. These key-value combinations can be used to query the OSM database and extract specific geographical objects of interest, covering services (schools, pharmacies, hospitals, etc.), shops (restaurants, banks, bakers, etc.), leisure activities (sport, music schools, sport playground, etc.) or cultural aspects (museum, archeological sites, historic buildings, etc.)



## USER COSTS

Free & open Access.

## DATA INFRASTRUCTURE



**Infrastructure needed:** GIS or software environment for statistical computing (R, Python, among others).

**Data repositories/storage needed:** The Geofabrik server makes available the OSM dataset for free. Several software packages are available to transform pbf format into more standard geographical format, by filtering the content of the OSM database regarding OSM tag of interest. Storing and documenting these extracts (database version, OSM id of geographical objects) may be considered for reproducibility purposes.



## DATA GOVERNANCE & MANAGEMENT

**Staff costs:** Data can be extracted using various tools linked to the OSM project. Its use requires a certain expertise in the world of geographic information and the structure of OSM data.

**Data analysis needed:** Managing OSM geometries variety: data can be contributed through points (nodes), polygons (ways) or relations (lines) depending on the tag of interest.

**Quality assurance:** It requires important consistency checks due to a lack of international harmonization, poor completeness, multiple database entries, errors and unclear standards. Its use also implies to identify the relevant key-values to catch all the possibilities of contribution referring to a specific service. Most of the studies concluded that OSM can efficiently complete authoritative or institutional data sources. It underlines also that POI completeness is very heterogenous, depending on the POI (banks or a pharmacies are significantly more complete than hairdressers) or the area of interest (rural areas are significantly less contributed than urban areas).

## DATA DOCUMENTATION AND/OR ASSISTANCE IN DATA USE



[OSM wiki](#); [OSM taginfo](#); [Geofabrik – OpenStreetMap Data Extracts](#); [Overpass turbo](#); [Introducing osmextract \(R\)](#); [Pyrosm \(Python\)](#).

### How to cite

© OpenStreetMap contributors



GRANULAR

## DATA FICHE

# Street-level imagery by Mapillary

<https://www.mapillary.com/>

Ian McCallum & Martin Hofer (IIASA)

### INDICATOR CLASS: Residential



#### Data

Class: images/video

Type: API

License: Creative Commons Share Alike (CC BY-SA)



#### Spatial

Extent: Global

Granularity: street-level

Coordinate System: World Mollweide (ESRI: 54009) and WGS 1984 (EPSG: 4326)

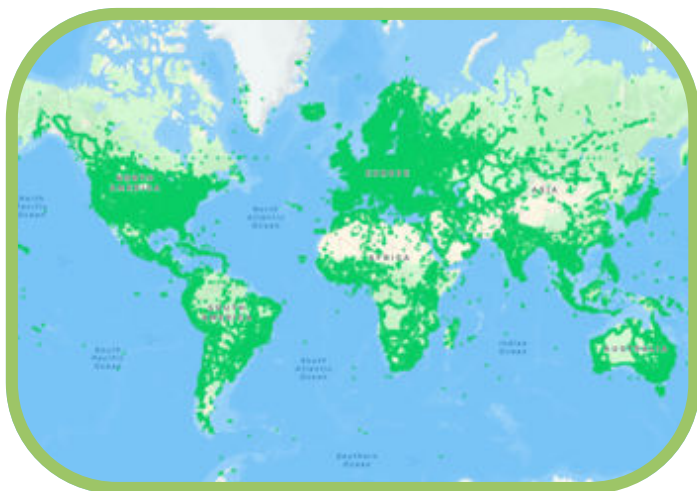


#### Temporal

Resolution: Daily

Extent: 2013 - present

Frequency: Random – based on when volunteers upload their data



### DESCRIPTION

Mapillary brings together a **global network of contributors** who want to make the world accessible to everyone by **visualizing the world and building better maps**.

Anyone can join and collect street-level images, using simple tools like smartphones or action cameras. With computer vision, we connect images across time and space to create **immersive street-level views** and **extract map data**.



## USER COSTS

Free & open Access.

## DATA INFRASTRUCTURE



Since joining Meta, Mapillary no longer charges for access to the Mapillary platform and Mapillary data. You are welcome to use the platform free of charge, as long as you comply with the Terms of Use. All imagery that you upload to Mapillary can be used by yourself in any form. Public imagery with a CC-BY SA licence is also available for use.

Plugins come for ArcGIS which is proprietary software. However, the imagery is also available for OSM. In addition, derived products are available in shapefile and GeoJSON formats for use in open software.



## DATA GOVERNANCE & MANAGEMENT

**Staff costs:** working with this data will require some expertise in handling GIS data and possibly computer vision. For a person with these skills, one person per month would likely be needed to investigate, extract and process the data.

**Data analysis needed:** Computer vision, data segmentation, and data extraction. This would be dependent on what information one wanted to extract.

**Quality assurance:** This would have to be done by the user on a case-by-case basis.

## DATA DOCUMENTATION AND/OR ASSISTANCE IN DATA USE



<https://help.mapillary.com/hc/en-us>

<https://forum.mapillary.com/>

### How to cite

Dataset: Lypez-Antequera, M., Gargallo, P., Hofinger, M., Bult, S.R., Kuang, Y., & Kotschieder, P. (2020). Mapillary Planet-Scale Depth Dataset. European Conference on Computer Vision.



# GRANULAR

## DATA FICHE

### Tourism by JRC

<https://tourism-dashboard.ec.europa.eu/?lng=en&ctx=tourism>

Lauri Niskanen (LUKE)

#### INDICATOR CLASS: Infrastructure, Tourism



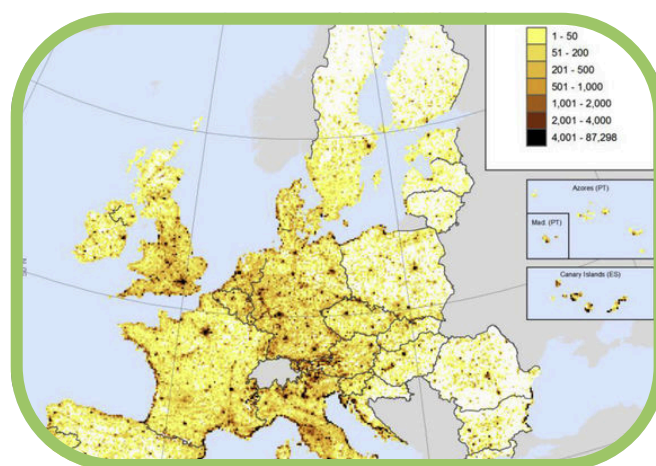
#### Data

Class: Webscraping

Type: Raster

License: The data in the Energy & Industry Geography Lab are free for public use

DOI: <https://doi.org/10.1016/j.tourman.2018.02.020>



#### Spatial

Resolution: 100m

Extent: EU

Coordinate System: Spherical/ Web Mercator

EPSG:3857



#### Temporal

Resolution: Extent: 2019-2022

Frequency: Annual

#### DESCRIPTION

**Tourism capacity and density** based on booking.com, TripAdvisor and Eurostat data. This descriptor measures the number of bed-places at tourism accommodation establishments available at a destination (country or region) in a given year. It provides an account of the absolute dimension of a tourism destination from the supply perspective. Higher values indicate a higher tourism supply. This descriptor is provided in the EU Tourism Dashboard as complementary information to characterise tourism supply. UNIT OF MEASURE: number of bed-places. RESOLUTION: NUTS0, NUTS2. COMPLETENESS: Missing countries for 2020: CH, CZ, DE, DK, HU, IE, LV, PL, SE, SI. DATA SOURCES: tour\_cap\_nuts2 [ESTAT], TripAdvisor.





## USER COSTS

Free & Open Access.

## DATA INFRASTRUCTURE



**Infrastructure needed:** None for viewing. Analyses require GIS or statistical software.

**Data repositories/storage needed:** Yes.



## DATA GOVERNANCE & MANAGEMENT

**Staff costs:** Basic GIS or statistical skills are needed, and more advanced analyses require a higher skill set.

**Data analysis needed:** No.

**Quality assurance:** None specified.

## DATA DOCUMENTATION AND/OR ASSISTANCE IN DATA USE



<http://data.europa.eu/89h/659a45dd-5bc2-4aaf-8bcd-bb337ba03f92>

### How to cite

Dataset: Ricardo Barranco (2022): UDP - Tourism capacity. European Commission, Joint Research Centre (JRC) [Dataset] PID: <http://data.europa.eu/89h/659a45dd-5bc2-4aaf-8bcd-bb337ba03f92>

Concept & Methodology:

Batista e Silva, F., Herrera, M. M., Rosina, K., Barranco, R. R., Freire, S., & Schiavina, M. (2018). Analysing spatiotemporal patterns of tourism in Europe at high-resolution with conventional and big data sources. *Tourism Management*, 68, 101-115. <https://doi.org/10.1016/j.tourman.2018.02.020>



GRANULAR

DATA FICHES

# MOBILITY

BENCHMARK OF PERFORMANCE AND COSTS ENABLING  
IDENTIFICATION OF VIABLE OPTIONS GOING FORWARD





GRANULAR

## DATA FICHE

# Human powered mobility by Strava Metro

<https://metro.strava.com/>

Ian McCallum & Martin Hofer (IIASA)

### INDICATOR CLASS: Residential



#### Data

Resolution: Class: vectors

Type: Shapefiles

License: <https://metro.strava.com/>



#### Spatial

Resolution: NA

Extent: Global

Granularity: street-level

Coordinate System: WGS 1984 (EPSG: 4326)



#### Temporal

Resolution: Annual

Extent: 2015 - present

Frequency: Updated monthly by users who agree to opt in to making their data public

### DESCRIPTION

The **Strava** dataset is the **largest collection of human-powered transport information** in the world. Metro aggregates, de-identifies and contextualizes this dataset to help make communities better for anyone on foot or a bike.

Strava works with urban planners, trail networks, city governments and safe-infrastructure advocates to understand mobility patterns, identify investment opportunities and evaluate the impact of infrastructure changes free of charge.



## USER COSTS

Free & open Access.

Since 2020, Metro is free to any organization that shares the mission to make cities better for cyclists and pedestrians. COVID-19 has accelerated the world's need for Metro – for smarter and more sustainable design of our cities, and for giving back to the communities that support millions of Strava athletes around the world. If you believe in human-powered transport and think you can make an impact, apply now. If your application is successful, you will be granted free access.

## DATA INFRASTRUCTURE



Information not available.



## DATA GOVERNANCE & MANAGEMENT

**Staff costs:** Using the online dashboard and UI would mean costs would be low, and anyone with basic internet skills could query and visualize the data.

**Data analysis needed:** If a user chooses to download the data, e.g. Shapefiles, this will then require GIS skills to work with and further analyze the data. This should however be fairly basic for any GIS analyst. Onin total would likely be enough to develop basic tools to work with and process the data.

**Quality assurance:** Strava is performing QA on their end, hence little effort is required for this from the user.

## DATA DOCUMENTATION AND/OR ASSISTANCE IN DATA USE



<https://metro.strava.com/> <https://www.strava.com/>

### How to cite

Dataset: [https://metro.strava.com](https://metro.strava.com/)





GRANULAR

DATA FICHES

# RECREATION

BENCHMARK OF PERFORMANCE AND COSTS ENABLING  
IDENTIFICATION OF VIABLE OPTIONS GOING FORWARD







# Potential quiet areas in Europe based upon the quietness suitability index

<https://www.eea.europa.eu/data-and-maps/figures/quietness-suitability-index-qi-and-1>

Antti Hiltunen (LUKE)

## INDICATOR CLASS: Health, Environment



### Data

Class: Model

Type: Vector + Raster

License: EEA standard re-use policy



### Spatial

Resolution: 100m, 3 arcsec

Extent: Europe

Granularity: Gridded

Coordinate System: ETRS89 (EPSG: 4258)



### Temporal

Resolution: ~Every 5 years

Extent: 2006, 2011, 2016

Frequency: >1 years

## DESCRIPTION

Quietness can be defined according to objective criteria (**noise levels**), which are measured by quantitative data, but also according to a subjective component linked to **perception**. Beyond noise exposure, quietness is described in the **Quietness Suitability Index (QSI)** as a combination of noise limit values (contour maps delivered under END requests) and land use and land cover elements that are perceived as positive and usually related to human cultural construction of naturalness.

QSI is composed of two main elements: (1) **Noise disturbance** as a result of distance to noise sources (objective criteria, quantitative data): threshold distances are determined considering noise levels determined by noise contours maps (END areas exposed to less than 55 dB Lden); (2) The **perceptive dimension of quietness by human beings** (subjective criteria, qualitative data): this dimension is related to the importance given to natural elements and landscape configuration. This dimension has been introduced in the QSI formula as a reclassification of the Corine Land Cover database based on the hemeroby index (Jalas, 1955; Blume and Sukopp, 1976).

To establish the distance values to the different noise sources, the noise contour maps for the main transport infrastructures and also the location of the major noise sources have been used as the main input information. These data have been provided by the Member States following the END requirements, accounting for main transport networks at the EU level and urban areas above 100,000 inhabitants. Text: EEA (2016). Quiet areas in Europe. Technical report No 14/2016. p. 12-13.



## USER COSTS

Free & Open Access.

## DATA INFRASTRUCTURE



**Infrastructure needed:** GIS/Remote sensing software.

**Data repositories/storage needed:** Yes, the files are quite large and require download to access.



## DATA GOVERNANCE & MANAGEMENT

**Data analysis needed:** No.

**Staff costs:** Basic GIS skills.

**Quality assurance:** Very high.

## DATA DOCUMENTATION AND/OR ASSISTANCE IN DATA USE



European Environment Agency (2016). Potential quiet areas in Europe, based upon Quietness Suitability Index (QSI). Retrieved from EEA's website: <http://data.europa.eu/88u/dataset/e9151c34-da65-48b9-a2ca-b9b835480812>.

EEA (2016). Quiet areas in Europe. Technical report No 14/2016: <https://www.eea.europa.eu/publications/quiet-areas-in-europe>.

### How to cite

Dataset: European Environment Agency (2016). Potential quiet areas in Europe, based upon Quietness Suitability Index (QSI). Retrieved from EEA's website: <http://data.europa.eu/88u/dataset/e9151c34-da65-48b9-a2ca-b9b835480812>  
Concept & methodology: EEA (2016). Quiet areas in Europe. Technical report No 14/2016 <https://www.eea.europa.eu/publications/quiet-areas-in-europe>



GRANULAR

DATA FICHES

# TRANSVERSAL

BENCHMARK OF PERFORMANCE AND COSTS ENABLING  
IDENTIFICATION OF VIABLE OPTIONS GOING FORWARD





# GRANULAR

## DATA FICHE

### Local Administrative Units geometries (LAU)

<https://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/administrative-units-statistical-units/lau>

Ronan Ysebaert (CNRS)

#### INDICATOR CLASS: Transversal



#### Data

Class: geometries

Type: vector (polygons)

License: Creative Commons Share Alike (CC BY-SA)



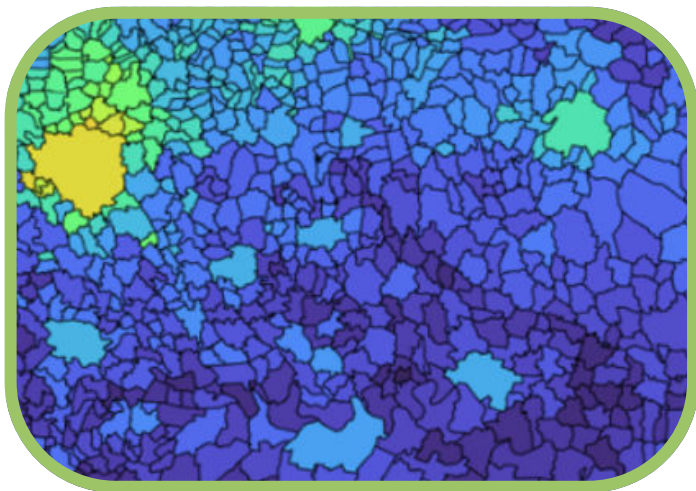
#### Spatial

Resolution: LAU

Extent: EU, EFTA, Candidate Countries (Serbia, Albania, North Macedonia) & UK (until the 2020 version)

Granularity: Municipalities & communes of the European Statistical System.

Coordinate System: ETRS89-extended / LAEA Europe (EPSG:3035), Web Mercator projection (EPSG:3857) and WGS 1984 (EPSG: 4326)



#### Temporal

Resolution: Annual

Extent: 2011-2021

Frequency: Annually since 2011

#### DESCRIPTION

Local Administrative Units (LAUs) are the **building blocks** of the Nomenclature of Territorial Units for Statistics (NUTS) and statistical regions. It covers municipalities and communes of the European Statistical System (ESS). Data is available annually and currently described by total resident population, where available. This is the core layer to fit with the national censuses of the European Union.



## USER COSTS

Free & open Access.

## DATA INFRASTRUCTURE



**Infrastructure needed:** GIS or software environment for statistical computing.

**Data repositories/storage needed:** No, data is moreover distributed via API: Several software API packages help to retrieve Eurostat – GISCO data such as giscoR in R or happygisco in Python: LAU units can be uploaded directly via computer programming.



## DATA GOVERNANCE & MANAGEMENT

**Staff costs:** Basic GIS skills.

**Data analysis needed:** No, just requires finding the appropriate LAU version to fit with data provided by Eurostat/National Statistical Institutes.

**Quality assurance:** Very High. Seamless geometries for all of Europe are a real added value to map results at the local scale. However, this territorial nomenclature is quite heterogeneous and highly subject to MAUP effects (35,000 units for France, 290 for Sweden). It must be considered when proposing indicators at this territorial scale.

## DATA DOCUMENTATION AND/OR ASSISTANCE IN DATA USE



[Eurostat, 2023, Local Administrative Units \(LAU\).](#)

[Eurostat, 2019, Merging statistics and geospatial information, Experiences and Observations from National Statistical Authorities](#)

### How to cite

Dataset: © EuroGeographics for the administrative boundaries





# GRANULAR

## DATA FICHE

### GEOSTAT 1km population grid

<https://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/grids>

Ronan Ysebaert (CNRS)

#### INDICATOR CLASS: Transversal



#### Data

Class: geometries

Type: vector (polygons, points), raster and tabular (centroids)

License: Creative Commons Share Alike (CC BY-SA)



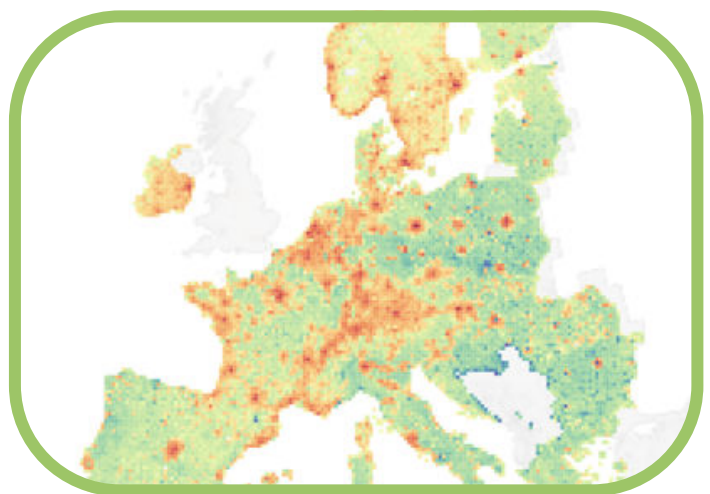
#### Spatial

Resolution: 1 km, also aggregated at 2-5-10-20-50-100 km

Extent: EU, Liechtenstein, Switzerland & Norway

Granularity: Grid cell

Coordinate System: ETRS89-extended / LAEA Europe (EPSG:3035), Web Mercator projection (EPSG:3857) and WGS 1984 (EPSG: 4326)



#### Temporal

Resolution: Total population for year YYYY.

Extent: 2006 – 2011 – 2018 - 2021

Frequency: Certainly more frequent in the future.

#### DESCRIPTION

The grid data consists of **statistics that are geographically referenced to a system of square grid cells in a common European reference grid net** with Cartesian coordinates. The use of a common European-level grid allows analyses to be undertaken for areas that cross national boundaries.

The net 1 km grid covering the territory of the EU contains **4 448 981 grid cells** and **1 824 619 populated cells**. In 2024, the 2021 population grid has been updated with new indicators: sex (males, females), age (under 15, 15 to 64, 65 and over), employed persons (as far as possible), place of birth (in the reporting country, in another EU country, outside the EU), usual residence 12 months before the census date (unchanged, within the reporting country, outside of the reporting country).



## USER COSTS

Free & open Access.

## DATA INFRASTRUCTURE



**Infrastructure needed:** GIS or software environment for statistical computing.

**Data repositories/storage needed:** No, data is moreover distributed via API: Several software API packages help to retrieve Eurostat – GISCO data such as giscoR in R or happygisco in Python: grid cells can be uploaded directly via computer programming.



## DATA GOVERNANCE & MANAGEMENT

**Staff costs:** GIS skills needed

**Data analysis needed:** Data handling requires a smart approach, in particular when manipulating vector data (computing time/cost).

**Quality assurance:** High in general, and growing over time. Three methodologies are used to attribute several inhabitants to each square kilometre cell: 1/ Aggregation method, using geo-referenced microdata (a method used by 25 Member States and 3 EFTA countries for the 2021 product). 2/ Disaggregation method is used in the absence of geocoded microdata, using statistical data for the lowest available administrative/territorial units in combination with auxiliary spatial data (also called the top-down approach). None of the Member States used this method for the 2021 population and housing census, but it was the case for previous grid products. 3 / Hybrid method that combines aggregation and disaggregation techniques and represents a compromise between accuracy and availability of data. It is used in France and Greece. Analysis using evolution over time must consider these aspects (methodologies evolving across EU countries).

## DATA DOCUMENTATION AND/OR ASSISTANCE IN DATA USE



[Eurostat, 2023, Statistics Explained, Population and housing census 2021, population grids](#)  
[Eurostat, 2023, Statistics Explained, the degree of urbanisation manual - Constructing a population grid](#)  
[European Forum for Geostatistics, 2012, ESSnet project GEOSTAT, Representing Census data in a European Population grid, Final Report GEOSTAT 1A – Representing Census data in a European population grid](#)

### How to cite

Dataset: © EuroGeographics for the administrative boundaries



# GRANULAR

## DATA FICHE

### NUTS geometries

<https://ec.europa.eu/eurostat/fr/web/gisco/geodata/reference-data/administrative-units-statistical-units/nuts>

Ronan Ysebaert (CNRS)

#### INDICATOR CLASS: Transversal



#### Data

Class: geometries

Type: vector (polygons, lines, points)

License: Creative Commons Share Alike (CC BY-SA)



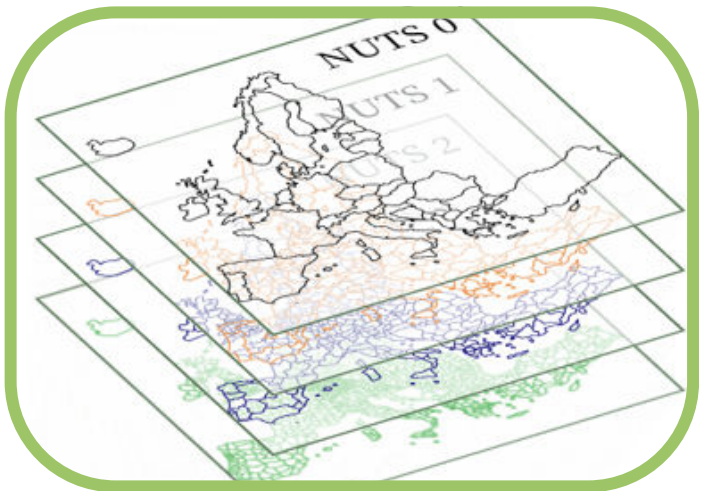
#### Spatial

Resolution: NUTS 0-1-2-3

Extent: EU, EFTA, Candidate Countries + the UK.

Granularity: 5 levels of detail (scales): 01M, 03M, 10M, 20M, 60M

Coordinate System: ETRS89-extended / LAEA Europe (EPSG:3035), Web Mercator projection (EPSG:3857) and WGS 1984 (EPSG: 4326)



#### Temporal

Resolution: Depending on territorial reforms

Extent: 2003, 2006, 2010, 2013, 2016, 2021

Frequency: every 3-4 years in general

#### DESCRIPTION

The Nomenclature of Territorial Units for Statistics (NUTS) is a **hierarchical system** divided into **3 levels**.

- NUTS 1: major socio-economic regions,
- NUTS 2: basic regions for the application of regional policies,
- NUTS 3: small regions for specific diagnoses.

The NUTS legislation is periodically amended; therefore multiple years are available for download. The NUTS geometries are therefore the basis for mapping statistical information made available by Eurostat.



## USER COSTS

Free & open Access.

## DATA INFRASTRUCTURE



**Infrastructure needed:** GIS or software environment for statistical computing.

**Data repositories/storage needed:** No, data is moreover distributed via API: Several software API packages help to retrieve Eurostat – GISCO data such as giscoR in R or happygisco in Python. NUTS units can be uploaded directly via computer programming.



## DATA GOVERNANCE & MANAGEMENT

**Staff costs:** Basic GIS skills needed.

**Data analysis needed:** No, it just requires finding the appropriate NUTS version to fit with Eurostat data.

**Quality assurance:** Very High.

## DATA DOCUMENTATION AND/OR ASSISTANCE IN DATA USE



<https://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/administrative-units-statistical-units/nuts>

### How to cite

Dataset: Administrative boundaries: © EuroGeographics, © TurkStat. Source: European Commission – Eurostat/GISCO